DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service

Food and Drug Administration
Rockville MD 20857

OCT 22 2004

Sharon W. Brown
Director, Drug Regulatory Affairs
Berlex Laboratories, Inc.
P. O. Box 1000
Montville, NJ 07045

Marie Kuker, R.Ph., RAC
Manager, Regulatory Affairs, US and International
3M Pharmaceuticals SBC
Bldg. 270-3A-08, 3M Center
St. Paul, MN 55144-1000

Docket No. 2002P-0029/CP1

Dear Ms. Brown and Ms. Kuker:

This letter responds to the citizen petition (Petition) dated January 16, 2002, and supplements dated November 27, 2002, and August 27, 2003, submitted by Berlex Laboratories, Inc., and 3M Pharmaceuticals (hereafter collectively referred to as Berlex). The Petition asks the Food and Drug Administration (FDA) to (1) change the therapeutic equivalence code of the estradiol transdermal system (ETS) manufactured by Mylan Technologies, Inc. (Mylan) from A-rated to B-rated, (2) change the labeling for the Mylan product to not allow placement of the patch on the buttock, (3) consider the Mylan product misbranded under sections 502(a), (f), and (j) of the Federal Food, Drug, and Cosmetic Act (the Act)[1], and (4) switch Mylan's application from a 505(j) to a 505(b)(2) submission. Petition at 1. In answering this petition, FDA has considered Mylan's comments dated April 17, 2003. As detailed below, Berlex relies upon data from a single study, and these data are inadequate to support any of its requests. Accordingly, the Petition is denied.

I. Summary of Statutory and Regulatory Basis for ANDA Approval

The Drug Price Competition and Patent Term Restoration Act of 1984 (the Hatch-Waxman Amendments) created section 505(j) of the Act, which established the current approval process for abbreviated new drug applications (ANDAs). The showing that must be made for an ANDA to be approved is different from that which is required in a new drug application (NDA). An NDA applicant must prove that the drug product is safe and effective. An ANDA applicant does not have to prove the safety and effectiveness of the drug product because an ANDA relies on FDA's previous finding that the reference listed drug is safe and effective. To rely on this finding, however, an ANDA applicant must demonstrate, among other things, that its generic[2] drug product is

[1] 21 U.S.C. 352 (a), (f), (j).

[2] For purposes of this response, the term "generic" refers to new drug products for which approval is sought in an ANDA submitted under section 505(j) of the Act (21 U.S.C. 355(j)).

02P-0029

PDN 1

bioequivalent to the reference listed drug. 21 U.S.C. 355(j)(2)(A)(iv). "Bioequivalence means the absence of a significant difference in the rate and extent to which the active ingredient or active moiety . . . becomes available at the site of drug action when administered at the same molar dose under similar conditions in an appropriately designed study." 21 CFR 320.1(e). The scientific premise underlying the Hatch-Waxman Amendments is that drug products that are bioequivalent and pharmaceutically equivalent[3] and, therefore, therapeutically equivalent, generally may be substituted for each other.[4]

A generic drug product is bioequivalent to the reference listed drug if

> the rate and extent of absorption of the drug do not show a significant difference from the rate and extent of absorption of the listed drug when administered at the same molar dose of the therapeutic ingredient under similar experimental conditions in either a single dose or multiple doses
> ...
>
> 21 U.S.C. 355(j)(8)(B)(i)); see also 21 CFR 320.1(e) and 320.23(b).

FDA regulations at 21 CFR part 320 establish acceptable methodologies for determining the bioequivalence of drug products. These include pharmacokinetic studies, pharmacodynamic studies, comparative clinical trials, and in vitro studies. The choice of study design to use is based on the ability of the design to compare the drug delivered by the two products at the particular site of action of the drug. The courts have expressly upheld FDA's regulatory implementation of the Act's bioequivalence requirements. See, e.g., *Schering Corp. v. FDA*, 51 F.3d 390 at 397-400 (3rd Cir. 1995); *Fisons Corp. v. Shalala*, 860 F. Supp. 859 (D.D.C. 1994).

II. Standard Bioequivalence Testing[5]

The standard bioequivalence (pharmacokinetic) study is conducted using a two-treatment crossover study design in a small number of volunteers, usually healthy,

---

[3] Pharmaceutically equivalent drug products have identical dosage forms and contain identical amounts of the identical active ingredient, and meet the identical compendial or other applicable standard of identity, strength, quality, and purity. They do not necessarily contain the same inactive ingredients and may also differ in characteristics such as shape, scoring, release mechanism and, within certain limits, labeling. See 21 CFR 320.1; FDA's *Approved Products with Therapeutic Equivalence Evaluations* (Orange Book), preface, p. vii.

[4] A generic drug that establishes bioequivalence as well as pharmaceutical equivalence is A-rated as therapeutically equivalent to the reference listed drug in the Orange Book. Drug products that the Agency does not currently consider therapeutically equivalent to other pharmaceutically equivalent products are B-rated.

[5] See Orange Book (preface) for further discussion of statistical criteria for bioequivalence and bioequivalence more generally.

normal adults. Single doses of the test and reference drug products are administered to each volunteer, and the blood, plasma, or serum levels of the drug are measured over time. The pharmacokinetic parameters characterizing the rate and extent of absorption are examined by statistical procedures. The pharmacokinetic parameters of interest are: (1) the resulting area under the curve when graphing plasma concentration over time (AUC), calculated to the last measured concentration ($AUC_{(0-t)}$) and extrapolated to infinity ($AUC_{(0-inf)}$), for extent of absorption; and (2) the maximum or peak drug concentration (Cmax) for rate of absorption.

The statistical methodology for analyzing these bioequivalence studies is called the "two one-sided test" procedure. Two situations are tested with this statistical methodology. The first of the two one-sided tests determines whether a generic (test) product, when substituted for a brand-name (reference) product, is significantly less bioavailable. The second of the two one-sided tests determines whether the reference product, when substituted for the test product, is significantly less bioavailable. Under FDA's longstanding policy, which was based on the opinions of medical experts and has been substantiated through many years of regulatory experience, a result showing a difference of greater than 20 percent between the test and reference products in either of the above tests is considered significant and, therefore, undesirable. Accordingly, the Agency has set a difference of 20 percent as a bound for acceptability. Numerically, these bounds are expressed as a limit for the first statistical test of 80 percent for the ratio of the average results for the test product to the average results for the reference product and, correspondingly, for the second test, a limit of 80 percent for the ratio of the reference-product average to the test-product average. By convention, all data are expressed as a ratio of the average response (AUC and Cmax) for the test product to the reference product, so the limit for the second statistical test is expressed as 125 percent (the reciprocal of 80 percent).

In practice, these statistical tests are carried out using an analysis of variance (ANOVA) procedure to calculate a 90 percent confidence interval for the values of both Cmax and AUC. The confidence interval for each of AUC and Cmax should fall entirely within the 80 percent to 125 percent boundaries described above. Because the mean of the study data lies in the center of the 90 percent confidence interval, the ratio of the means of the data for the test product to the reference product is usually close to 100 percent (a test/reference ratio of 1).

III. Background Concerning Climara and the Mylan ETS

The reference listed drug for Mylan's ETS is Climara, which is jointly manufactured and marketed by Berlex and 3M pursuant to NDA 20-375. Climara is used for hormone replacement therapy. It is indicated for the relief of symptoms associated with menopause, including vasomotor symptoms (hot flushes) and vulval and vaginal atrophy. It is also used to treat low estrogen levels caused by hypogonadism, castration, or primary ovarian failure. In addition, Climara is indicated for treatment of abnormal uterine bleeding caused by hormonal imbalance in the absence of organic pathology when associated with a hypoplastic or atrophic endometrium.

Climara is available in the following four strengths: 0.05 milligram (mg)/24 hours (hr), 0.1 mg/24 hr, 0.075 mg/24 hr, and 0.025 mg/24 hr. Each of these four strengths is designated as a reference listed drug in FDA's Orange Book.

Mylan manufactures two ETS products: one in a 0.05 mg strength that is AB-rated (as therapeutically equivalent) to Climara's 0.05 mg strength and one in a 0.1 mg strength that is AB-rated to Climara's 0.1 mg strength.

Climara was originally approved in 1994 for application to the abdomen only. In 1997, Berlex submitted a supplement to its NDA containing a bioequivalence study comparing the bioavailability of Climara when applied to the buttock and when applied to the lower abdomen. The study did not show that the administration at the buttock site was bioequivalent to administration at the approved abdomen site. When Climara was applied to the buttock site, Cmax and AUC were 25 percent and 17 percent higher, respectively, than when Climara was applied to the abdomen site. Nonetheless, the Agency concluded that the plasma levels of estradiol fell within the expected therapeutic range when Climara was applied to the buttock. Based on this conclusion, the Agency approved the buttock application site for Climara on April 11, 1997. The approved labeling for Climara states that the product can be applied either to the abdomen or to the buttock.

Berlex submitted a citizen petition and petition for stay of action (Docket No. 98P-0434/CP1 and PSA1) dated June 12, 1998, asking FDA not to approve a generic version of Climara unless certain scientific, medical, regulatory, and legal criteria were met. Among these criteria, Berlex requested that an ANDA applicant for a generic version of Climara show bioequivalence at both application sites because the bioavailability of estradiol is different on the abdomen site than on the buttock site. On March 17, 2000, FDA denied this request and denied the citizen petition overall. The Agency's response stated in pertinent part:

> If a bioequivalent product is placed on the same site as the reference listed drug product, it will yield equivalent plasma concentrations. If these same two products are placed on an alternate site on the body, different plasma concentrations may result, but the products will still yield equivalent plasma concentrations.
>
> Petition response at 21.

FDA approved Mylan's ANDA in February 2000 with labeling that allows use of the product at either the abdomen or the buttock site.

IV. Discussion of Berlex's Arguments

In December 2001, Berlex submitted to the Agency a bioequivalence study it had performed comparing estradiol delivery from the Mylan and Climara transdermal systems on the buttock site in 40 healthy postmenopausal women. The study was an open-label, randomized, 3-period, crossover study using a dose of 0.1 mg per day. Based on this study, Berlex argues that the Mylan ETS applied to the buttock is not therapeutically equivalent to Climara, and therefore, cannot be labeled for this use. We disagree. As explained below, the data Berlex has submitted do not provide an adequate basis for challenging the Agency's conclusion that the Mylan ETS is therapeutically equivalent to Climara when applied at the buttock.

A. Berlex Data Inadequate to Assess Equivalence of Cmax

Berlex states that its study showed that the Cmax of the Mylan ETS averaged about 16 percent higher than the Climara Cmax, and the 90 percent confidence interval for the Mylan ETS ranged from 107 percent to 126 percent. Berlex asserts that this difference exceeds the Agency's bioequivalence limit of 125 percent. Berlex states that patients who are switched from Climara to the Mylan ETS will receive almost 14 percent more estradiol over the 7-day dosing period than with Climara. Berlex concludes that the Mylan ETS applied to the buttock is not bioequivalent to Climara. Petition at 4.

Berlex's data do not support this conclusion because the study was not adequately powered, i.e., included too few subjects, to be able to determine whether the products are bioequivalent at the site. Statistical power generally means the probability that an effect can be detected if it, in fact, exists. In the context of a bioequivalence study, power means the probability that the study will be able to show whether the generic product is, in fact, bioequivalent to the innovator product. Accordingly, if a bioequivalence study is underpowered (has too few subjects), the study may not result in a finding of bioequivalence even though the two products are actually bioequivalent.

In a crossover design study (used in both the Berlex and Mylan bioequivalence studies), the amount of power for a given sample size is determined by two factors: intrasubject variability (variability in the rate and extent of absorption in the same individual following administration of the same drug product at different times) and the expected difference in pharmacokinetic values between the two products. The greater the intra-subject variability and/or difference in pharmacokinetic values, the larger the sample size needed to power the study adequately.

Berlex assumed an intrasubject variability for both Climara and the Mylan ETS of 25 percent, and FDA is not challenging this assumption. However, Berlex's assumption about the expected difference in the pharmacokinetic values of the products was incorrect. Berlex assumed that the difference in AUC and Cmax values for Climara and the Mylan ETS would be plus or minus 5 percent. Berlex's study results suggest that the difference in AUC and Cmax between Climara and the Mylan ETS is not 5 percent, however, but at least 10 percent.

To obtain similar power, the sample size needed with a difference of ten percent in pharmacokinetic (AUC and Cmax) values is significantly greater than would be needed if the difference were five percent (the assumption made by Berlex).[6] Because the differences in AUC and Cmax were larger than Berlex had assumed when deciding how to power its study, the sample size of 39 subjects in the Berlex study was not large enough (had inadequate power) to assess the bioequivalence of the two products at the buttock site.[7] (The study originally enrolled 42 subjects, but 2 subjects dropped out and Berlex excluded 1 subject.)

Berlex asks FDA to reverse the Agency's determination that the two products are bioequivalent at the buttock site, a determination based upon the Agency's well-established practices and experience in making bioequivalence determinations. Yet, with a sample size of 39 subjects, the Berlex study has less power than a bioequivalence study submitted by an ANDA applicant generally would.[8] The Agency sees no reason to

---

[6] As the ratio between test and reference products deviates from one, the number of subjects needed generally increases in a non-linear fashion. For example, for 90 percent power (90 percent confidence in the results), assuming an intrasubject variability of 25 percent and no differences between the two formulations, 28 subjects would be needed to show bioequivalence. If the difference between formulations were 5 percent, 36 subjects would be needed to show bioequivalence, and 66 subjects would be needed if the difference between formulations were 10 percent. E. Diletti, et al., 1992, "Sample size determination for bioequivalence assessment by means of confidence intervals," *Int J Clin Pharmacol Ther Toxicol*, 30 Suppl 1:S51-8.

[7] Because the Berlex study suggested that the difference in AUC and Cmax between Climara and the Mylan ETS was at least 10 percent, Berlex's claim (on page 4 of its August 27, 2003, supplement) that its study was actually sufficiently powered, if the difference between the products had been 8 percent, is irrelevant.

[8] Berlex asserts that "if one were to use the Mylan study results . . . to help calculate the appropriate sample size for the Berlex study, one would have selected 32 subjects. This in fact is in excellent agreement with the sample size calculation for the Berlex study." August 27, 2003, supplement at 10. This argument fails to take into account, however, that the use of a replicate design by Mylan provided power equivalent to that of a non-replicate crossover design study with approximately twice the sample size.

Berlex also states: "[W]e do not accept this [Mylan's] contention, that simply passing the confidence interval criteria, confers therapeutic equivalence on the product." *Id.* The agency agrees that a demonstration of bioequivalence alone is not sufficient to establish therapeutic equivalence. To be rated therapeutically equivalent, the generic product must also be pharmaceutically equivalent. See footnote 3 and related text, p. 2.

Berlex goes on to suggest that a finding of bioequivalence could be obtained by using an excessive number of subjects (450), but that, in such a case, the products found to be bioequivalent should not necessarily be considered therapeutically equivalent. Berlex's concern is not relevant here, however. The issue in this instance is not the use of an excessive number of subjects, but Berlex's failure to use a sufficient number of subjects. Mylan used a sufficient number of subjects while Berlex did not.

reverse its position on the basis of such uncertain data.[9] To the contrary, the Agency generally would not consider it appropriate to reverse a prior determination of bioequivalence except on the basis of data that are more persuasive than the evidence upon which the initial determination was made.

In sum, because Berlex's study was not sufficiently powered to determine bioequivalence, its results do not undermine FDA's conclusion (based on scientific experience and expertise) that two transdermal products shown to be bioequivalent at one site are also bioequivalent to each other at an alternate site.

B. Mylan Blood Sampling Times Adequate to Demonstrate Tmax

Berlex asserts that the blood sampling times used in the bioequivalence study Mylan performed for approval of its ANDA were not adequate to assess bioequivalence. Petition at 5. In Mylan's study, samples were collected at 6, 12, 24, 48, 72, 96, 120, 144, and 168 hours after application of the patch. Berlex states that this sampling schedule incorrectly assumes that peak absorption occurs by 24 hours. Berlex states that "[n]umerous figures in the Climara product labeling, however, show a characteristic peak concentration of estradiol near 36 hours." Petition at 5. Berlex claims that the Mylan bioequivalence study cannot show whether the Mylan ETS has this 36-hour peak because no blood samples were taken between 24 and 48 hours.

However, data submitted under NDA 20-375 (Climara) showed that the application of different Climara patches (0.05 mg/day, 0.06 mg/day, and 0.1 mg/day) to the abdomen site all produced estradiol serum profiles with two peaks. The first peak occurs at ≤ 24 hours, while the second peak appears at ≤ 33 hours. Large intersubject variability was also reported for the estimates of Tmax. For example, a standard deviation of up to 11 hours was observed. The double peaks and the large intersubject variability do not support a conclusion that there is a precise, characteristic peak for Tmax, whether at 36 hours or at some other time between 24 and 48 hours. In short, Berlex offers no evidence to suggest that the fact that Mylan did not include a sampling point between 24 and 48 hours would have significantly affected the determinations of AUC and Cmax upon which the finding of bioequivalence was based.

Rather than discuss the actual underlying data, Berlex points to figures in Climara's labeling to support the claim of a Tmax value near 36 hours. However, figure 3 in the labeling shows only the mean data, suggesting that the product has consistent, precise, characteristic peaks, when, as explained above, Tmax actually is characterized by substantial variability. Further, this figure shows that Cmax occurs at 24 hours following application of the 6.5 cm^2 (0.025 mg/day) patch and that, because of double peaks of nearly equal magnitude, Tmax could be either 24 or 36 hours for the 12.5 cm^2 (0.05 mg/day) patch. Figure 4 compares the mean serum profiles following application

[9] In addition, FDA believes that the inclusion of an unrelated third arm in the study, to test a new formulation of Climara, may have confounded the study results by increasing the potential for a carryover effect.

of the 25 cm^2 (0.1 mg/day) patch (the dosage strength compared to Mylan's ETS in the bioequivalence study Berlex relies on in its petition) to the abdomen and to the buttock. While this figure shows a Tmax of 24 hours for the buttock and 36 hours for the abdomen, it also fails to fully reflect the substantial variability in Tmax observed in the underlying data upon which the figure relies. Moreover, figure 4, as well as figure 3, presents only an approximation of the precise estimation of Tmax obtained from the statistical analysis of the data itself.

In sum, Berlex has not adequately supported its argument that the sampling times in Mylan's bioequivalence study were inadequate.

C. Berlex Skin Adhesion Data Inadequate to Assess Relative Skin Adhesion

Berlex states that to be considered bioequivalent to Climara, the Mylan product must possess skin adhesion characteristics similar to those of Climara. Berlex reports that in its study, patch liftoff or falloff occurred in 59 percent of the applications of the Mylan ETS, compared with 18 percent of the Climara applications. According to Berlex, the median liftoff time for the Mylan product was 35.5 hours after application (23 occurrences), while the median liftoff time for Climara was 119 hours (6 occurrences).[10]

Berlex's conclusions about the adhesion of the Mylan ETS are not persuasive because Berlex did not follow FDA-recommended procedures for collecting and analyzing skin adhesion data or other procedures appropriate to generate precise or reliable results. Berlex proposed skin adhesion evaluation procedures in its initial, June 12, 1998, petition. In its March 17, 2000, response, the Agency stated that it would ask generic applicants to apply a test methodology similar to that proposed by Berlex, including an adhesion scoring system as part of a periodic visual assessment of adhesion. The Agency also agreed with Berlex's conclusion that the use of overlays should not be permitted, stating that withdrawing subjects from the study would be more scientifically sound.

Consistent with these Agency-endorsed recommendations, Mylan assessed skin adhesion in the same studies that assessed bioequivalence and skin irritation. Taping of patches was not permitted, and patch adhesion was evaluated by study evaluators at 6, 12, 24, 48, 72, 96, 120, and 144 hours after application. Patch adhesion was rated by a score from 0 to 7. A score of 7 meant that the patch remained completely attached; a score of 0 meant that the patch had come off. Scores between 0 and 7 indicated varying degrees of patch liftoff.

---

[10] Berlex states: "This disparity results in a significant pharmacoeconomic impact due to the increased costs associated with taping and replacing the less adherent Mylan ETS patches." Petition at 6; see Supplement at 2. However, as stated in our March 17, 2000, response, the Agency would not preclude an ANDA applicant from providing optional overlays with its marketed product for patient convenience.

By contrast, Berlex did not follow its own recommendations that the Agency had supported. Specifically, its protocol failed to include a systematic, scored methodology for visual assessment. Further, Berlex permitted patches that failed to adhere to be taped and left the decision to tape entirely to the study subjects.[11] In addition, study participants were allowed to determine when patch liftoff had occurred. In sum, Berlex's data are not reliable and do not constitute persuasive evidence that the Mylan ETS is inequivalent to Climara in skin adhesion.

However, patch adhesion is a product quality issue apart from its significance for bioequivalence. After approval, FDA continues to monitor and investigate reports of adhesion failure and requires sponsors to take appropriate action to address any problems and ensure product quality. To limit the risk to patients, drug product labeling recommends that "[i]n the event that a system should fall off, a new system should be applied for the remainder of the 7-day dosing interval."

V. Conclusion

Berlex's study was underpowered and, therefore, could not generate adequate data to assess the bioequivalence of the Mylan ETS to Climara at the buttock site. Berlex's data also do not show a failure to identify Tmax for the Mylan ETS and, therefore, do not show that Cmax was inaccurately determined as a result of such a failure. Berlex's data on skin adhesion are imprecise and unreliable and, therefore, do not offer an adequate basis for questioning the skin adhesion of the Mylan ETS. In short, the Berlex data offer no adequate basis for concluding that the Mylan ETS was not properly approved under an ANDA for use at both the abdomen and the buttock and properly determined to be therapeutically equivalent to Climara. Therefore, there is no basis upon which to grant Berlex's request for Mylan's application to be switched from an ANDA to a 505(b)(2) application, to change the therapeutic code of the Mylan ETS, to amend its labeling, or to render the product misbranded. Accordingly, the Petition is denied.

Sincerely,

Steven K. Galson, M.D., M.P.H.
Acting Director
Center for Drug Evaluation and Research

[11] Harrison, L. and D. Harari, 2002, "An evaluation of bioequivalence of two 7-day 17ß-estradiol transdermal delivery systems by anatomical site," *J Clin Pharmacol*, 42:1134-1141 at 1135.